UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR		ATTORNEY DOCKET NO.
08/970,889	11/14/97	BERGEN	J	SAR12228

WM51/1026

THOMASON MOSER & PATTERSON
2-40 BRIDGE AVENUE
PO BOX 8160
RED BANK NJ 07701-5300

EXAMINER	
PADMANABHAN, M	

ART UNIT	PAPER NUMBER
2671	16

DATE MAILED: 10/26/00

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner of Patents and Trademarks

PTO-90C (Rev. 2/95)

U.S. G.P.O. 2000 ; 465-188/25266

File Copy

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
ASSISTANT SECRETARY AND COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Paper No. 16

Application Number: 08/970889
Filing Date: 11/14/1997
Appellant(s): Bergen et al

MAILED
OCT 26 2000
Group 2700

Eamon J. Wall
For Appellant

EXAMINER'S ANSWER

The Group and/or Art Unit location of your application in the PTO has changed. To aid in correlating any papers for this application, all further correspondence regarding this application should be directed to Group Art Unit 2671.

This is in response to appellant's brief on appeal filed 9/6/2000.

(1) *Real Party in Interest*

A statement identifying the real party in interest is contained in the brief.

(2) *Related Appeals and Interferences*

A statement identifying the related appeals and interferences which will directly affect or be directly affected by or have a bearing on the decision in the pending appeal is contained in the brief.

(3) Status of Claims

The statement of the status of the claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Invention

The summary of invention contained in the brief is correct.

(6) Issues

The appellant's statement of the issues in the brief is correct.

(7) Grouping of Claims

Appellant's brief includes a statement that claims 1-3, 11, and 21-23 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

Appellant's brief includes a statement that claims 4 and 24 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

Appellant's brief includes a statement that claims 5-8 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

Appellant's brief includes a statement that claims 9-10 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

Appellant's brief includes a statement that claims 13-14 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

Appellant's brief includes a statement that claims 17-20 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

Appellant's brief includes a statement that claims 25-26 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

(8) Claims Appealed

The copy of the appealed claims contained in the Appendix to the brief is correct.

(9) Prior Art of Record

The following is a listing of the prior art of record relied upon in the rejection of claims under appeal.

5,706,417	Adelson	1/6/1999
5,751,286	Barber et al	5/12/1998
5,821,945	Yeo et al	10/13/1998
5,635,982	Zhang	6/3/1997

Shibata et al. ("Content-Based structuring of video information": 0-8186-7436-9/96, 1996 IEEE).

Jaillon et al. ("Image Mosaicing Applied to Three-Dimensional Surfaces": 1051-4651/94 - 1994 IEEE)

(10) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claims 1-11, 13-14, and 17-26 are rejected under 35 U.S.C. 103(a). This rejection is set forth in prior Office action, Paper No. 8, and is repeated here for reference.

Claim Rejections - 35 USC § 103

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Claims 1-3, 11, and 21-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al. ("Content-Based structuring of video information": 0-8186-7436-9/96, 1996 IEEE).

Claim 1 lays claim to a method of representing video information comprising the steps of segmenting a video stream into scenes, each scene into frames including a key frame, and also dividing scenes into at least one background and at least one foreground layer using intra-scene motion analysis, and storing content-related appearance attributes or mosaic representations in a database.

Claim 2 adds to claim 1 the step of computing, and storing content-related appearance attributes for the background and foreground layers.

Adelson teaches that a layer exists for each object, set of objects, or portion of an object in the image having a motion vector significantly different from any other object in the image (Col.2: lines 45-47). He also teaches combining the foreground and background images to produce a video image (Col.2: lines 15-21; Col.6: lines 50-55). Adelson also teaches content related appearance attributes for each layer with the use of intensity map, attenuation map, velocity map, and delta map (Col. 2: lines 50-67), and implicitly teaches storing these attributes in a database. While Adelson does not explicitly teach intra-scene motion, Adelson does teach a sequence of frames in Fig.4, wherein the foreground baseball object is defined by the frames e, f, h, I, j, and k, wherein the intra-scene motion analysis (since a scene may comprise of at least one frame, the intra-frame motion analysis, in this case, is equivalent to intra-scene motion analysis) is used to generate these foreground frames, and the background layer of frame d is formed by mosaicing the occluded region of the baseball in frame a with a similar region of the non-occluded background region. Furthermore, Adelson teaches using cumulative information of each frame to construct a lattice comprising the larger scene (mosaicing background), and also teaches displaying any portion of a scene that is desired to show, and also teaches extrapolating motion from other frames (Col.14, Figs.7A & 7B), implicitly teaching intra-scene motion analysis since each scene comprises of at least one frame, and conversely, each

frame comprises a portion of the scene. Adelson does not teach segmenting a video stream into scenes, and scenes into frames including a key frame. Yeo discloses dividing video sequence into equal length segments, denoting the first frame of each segment as its key frame (Col.1: lines 34-38), and also teaches classifying a long video sequence into story units (Col.1: lines 47-50) based on content, using temporal segmentation of video based on scene change detection. Yeo's keyframes teach the concept of detecting scene transitions between video sequences. Adelson teaches background mosaicing, and also teaches a foreground object that moves over the stationary background mosaic, and also teaches displaying portions of the scene that is desired to show, as explained above, and also teaches that while slow movements can be encoded as warps of a single layer, faster moving objects should be split into different layers (Col.15). Shibata teaches segmenting a video sequence, with individual video frames being the smallest unit of any segment. He also teaches the use of a basic segment which is a collection of video frames having the same vector expressions, assuming a collection of basic segments as the initial layer, and creating new layers by adding a segment to the previously processed layer, thus teaching a method for providing background mosaic, and intra-scene motion analysis. It would have been obvious to use intra-scene motion analysis to split the video information into layers, and make the static background mosaic of Adelson the keyframe, since this would provide a more efficient encoding of images.

Claim 3 adds to claim 2 the steps of storing the scenes in a mass storage unit, and retrieving scenes associated with an attribute.

Adelson teaches the use of video tape player, laser disc player as data source for image pixel data (Col.4: lines 2-7, 16-20). This implicitly teaches using mass storage unit to store data representing the scenes. Adelson also teaches having various maps for the various attributes (Col.2: lines 55-67; Col.5: lines 9-17), and retrieving data easily to reconstruct an image, based on the required image (Col.6: lines 30-47).

Claim 11 adds to claim 1 the steps of storing ancillary information related to layers or frames.

Adelson teaches the use of optional maps, including a contrast change map and a blur map for each layer (Col.3: lines 6-14).

Claim 21 is a claim for a computer readable medium that implements the method as claimed in claim 1 and hence is rejected for the same reasons.

Claim 22 is a claim for a computer readable medium that implements the method as claimed in claim 2 and hence is rejected for the same reasons.

Claim 23 is a claim for a computer readable medium that implements the method as claimed in claim 3 and hence is rejected for the same reasons.

Claims 4 and 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945) and Shibata et al, as applied to claims 1 and 22 respectively, and further in view of Jaillon et al. ("Image Mosaicing Applied to Three-Dimensional Surfaces": Jaillon et al.; 1051-4651/94 - 1994 IEEE).

Claim 4 adds to claim 1 the limitation that the mosaic representation is one of a two dimensional, a three dimensional, and a network of mosaics.

Jaillon teaches aligning and combining images or other mosaics to form a mosaic. Hence it would be obvious to one skilled in the art at the time the invention was made to combine various layers/images to generate a mosaic representation as this will provide the user greater flexibility in altering the image scene to suit their needs.

Claim 24 is a claim for a computer readable medium that implements the method as claimed in claim 4 and hence is rejected for the same reasons.

Claims 5-8 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al, as applied to claim 2, and further in view of Jaillon et al. ("Image Mosaicing Applied to Three-Dimensional Surfaces": Jaillon et al.; 1051-4651/94 - 1994 IEEE).

Claim 5 adds to claim 2 the steps of generating an image pyramid for a layer, filtering such that each subband is associated with feature maps, and integrating feature maps to produce

attribute pyramid subbands, which comprise content-based appearance attribute subband associated with a corresponding image pyramid subband.

Adelson discloses the use of subbands to encode images (Col.1: lines 20-24). Adelson also teaches the feature maps associated with each layer (Col.2: lines 55-67; Col.5: lines 9-17), and integrating the feature maps to reconstruct an image (Col.6: lines 30-47). Adelson and Yeo fail to teach image pyramids. Jaillon teaches the use of image pyramid framework in the alignment process, and converting the input image and the mosaic into Laplacian image pyramids, and applying the alignment to all levels within the respective pyramids. Hence it would be obvious to one skilled in the art at the time the invention was made to use the image pyramid in each layer in order to achieve better alignment and reproduction of the image.

Claim 6 adds to claim 5 the limitation that the attribute comprises at least one of luminance, chrominance, and texture.

Adelson discloses the use of intensity map, depth map, blur map, contract change map (Col.2: lines 55-67; Col.5: lines 9-17).

Claim 7 adds to claim 5 the step of rectifying the feature maps associated with each subband.

Adelson discloses the use of delta map, which is essentially an additive error map, which provides correction data for any changes in the image over time which can not be accounted for by the other maps.

Claim 8 adds to claim 5 the step of collapsing the attribute pyramid subbands to produce a content-based appearance attribute.

Yeo teaches that the lower levels of the hierarchy can be based on visual cues, while the upper levels allow criteria that reflect semantic information associated (Col.5: lines 48-52), the nodes capturing the contents of a video, while the edges capture its structure. Yeo also teaches a tree hierarchy that permits the user to have a coarse-to-fine view of the entire video sequences based on the level of the nodes (Col.4: lines 30-35), the nodes capturing the core contents of the video while the edges capture its structure (Col.5: lines 40-43). Hence it would be obvious to one skilled in the art at the time the invention was made to collapse the attribute pyramid subbands to produce a content-based appearance attribute since this will offer a browsing structure that closely resembles human perception and understanding.

Claims 9-10 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al., as applied to claim 2, and further in view of Barber et al. (U.S. Patent 5,751,286).

Claim 9 adds to claim 2 the steps of receiving a request matching a desired content-related appearance attribute, and retrieving at least one layer matching the request.

Adelson teaches a method of retrieving data representing layers, each layer comprising a series of maps, to reconstruct an image. Barber teaches a method of building a visual query by image content, and retrieving database images with features that correspond to the selected image

characteristics (Col.2: line 64 - Col.3: line 8). Hence it would be obvious to one skilled in the art at the time the invention was made to query the database by content-related appearance attribute, and retrieve layers that match the attribute, in order to reconstruct the image as desired, as such an approach will save database storage requirements.

Claim 10 adds to claim 9 the steps of identifying a query type as being one of luminance, chrominance, and texture type, and a query specification as being a desired property of the query type, and selecting a filter type and calculating the appearance attribute based on filter type and desired property.

Barber discloses a query construction interface with a hierarchical selection windows for each of image color, shapes, textures, category, which may include keywords, text or conditions (Col.3: lines 22-34). Barber also teaches filtering the masks in the current image by the category code, establishing the set of masks that will be analyzed with respect to the image characteristic values (Col.12: lines 1-5). Barber also teaches computing positional feature score that compares the area's similarity to the image areas (Col.14: lines 40-60). Hence it would be obvious to one skilled in the art at the time the invention was made to use a query type to chose the parameter, and specification to specify a desired property for the parameter, as this would facilitate retrieving only the layers that match the selection criteria, and hence would increase the speed of rendering the image.

Claims 13-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945) and Shibata et al., as applied to claim 1, and further in view of Zhang et al. (U.S. Patent 5,635,982).

Claim 13 adds to claim 1 the steps of generating a descriptor vector, and generating a scene cut indicium in response to calculated differences between descriptive vectors of successive frames exceeding a threshold.

Adelson teaches generating an intensity map, an attenuation map, a velocity map, and a delta map for each layer. Zhang teaches calculating the differences between consecutive video frames based on the selected difference metric, and defining a cut if the values exceed a threshold value (Col.7: lines 1-10; Col.8: lines 5-15). Hence it would be obvious to one skilled in the art at the time the invention was made to generate a scene cut if the calculated differences between descriptive vectors exceeded a threshold value, as this would minimize the calculation needed to detect scene cuts.

Claim 14 adds to claim 1 the steps of generating a descriptor vector and a threshold for it in the first pass, and calculating the difference between the frames and generating a scene cut indicium in the second pass, if the difference exceeds the threshold value.

Zhang teaches a multi-pass approach, wherein the prospective segment boundaries are determined in the first pass, by comparing against a threshold value. This implies the use of a descriptor vector to define a frame, such that they can be compared against a threshold value.

Zhang teaches using the second pass to locate all boundaries (scene cuts). Zhang also teaches using the multi-pass approach to apply different difference metrics in different passes (Col.6: lines 20-64), and teaches defining cuts based on the differences in the difference metrics (Col.8: lines 5-15). Hence it would be obvious to one skilled in the art at the time the invention was made to use two passes as described in this claim to compute the attribute value, as this would provide more accurate values for the attribute.

Claims 17-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Barber et al. (U.S. Patent 5,751,286) in view of Yeo et al. (U.S. Patent 5,821,945) and Shibata et al.

Claim 17 claims a method for browsing a video program comprising a plurality of scenes that contain frame(s), comprising the steps of providing a database comprising attribute information, formulating a query utilizing the attribute information, and searching and retrieving video frames that substantially match the query criterion.

Barber teaches a query facility which builds a visual query by image content, and also teaches a query engine that interprets the query, and returns database images with features that correspond to the selected criteria (Col.2: line 64 - Col.3: line 8). Barber does not teach the notion of a representative video frame for a video scene. Yeo discloses a method for content-based video browsing, containing a video database, and sets of key frames that have associated attributes, the key frames representing the long sequence of related shots (Col.2: lines 35-45).

Yeo also teaches the use of Rframes (representative frames) to organize the visual contents of the video clips (Col.1: lines 30-65). Shibata teaches segmenting a video sequence, with individual video frames being the smallest unit of any segment. He also teaches the use of a basic segment which is a collection of video frames having the same vector expressions, assuming a collection of basic segments as the initial layer, and creating new layers by adding a segment to the previously processed layer, thus teaching a method for providing background mosaic, and intra-scene motion analysis. Hence it would be obvious to one skilled in the art at the time the invention was made to build a query to retrieve the representative frames, as this would be a faster way to identify areas of interest before retrieving all the related frames.

Claim 18 adds to claim 17 the steps of selecting a query type, query specification, and computing a multi-dimensional feature vector.

Barber teaches query specification for image characteristics (query type) (Col.13: lines 44-53). Barber also teaches calculating a positional feature score combining features and positional similarity for each of the areas selected in the query (Col.15: lines 40-61).

Claim 19 adds to claim 18 the limitation of selecting a query specification by identifying a portion of the displayed image, and the feature vector is calculated based on query type and the identified image portion.

Barber teaches specification in a query of image characteristics that occur in some area or areas of the image (Col.13: lines 45-52).

Claim 20 adds to claim 19 the steps of formatting and transmitting the identified video frames.

Barber teaches returning the images with the best scores in response to a query (Col.14: lines 65-67).

Claims 25 - 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al, as applied to claim 22, and further in view of Jaillon et al. ("Image Mosaicing Applied to Three-Dimensional Surfaces": Jaillon et al.; 1051-4651/94 - 1994 IEEE).

Claim 25 is a claim for a computer readable medium that implements the method as claimed in claim 5 and hence is rejected for the same reasons.

Claim 26 is a claim for a computer readable medium that implements the method as claimed in claim 6 and hence is rejected for the same reasons.

(11) Response to Argument

Examiner disagrees with the appellants arguments that Adelson does not teach forming a background mosaic image, and Yeo does not teach mosaicing of background layers to form a

key frame. Adelson teaches a sequence of frames in Fig.4, wherein the foreground baseball object is defined by the frames e, f, h, I, j, and k, wherein the intra-scene motion analysis (since a scene may comprise of at least one frame, the intra-frame motion analysis, in this case, is equivalent to intra-scene motion analysis) is used to generate these foreground frames, and the background layer of frame d is formed by mosaicing the occluded region of the baseball in frame a with a similar region of the non-occluded background region. Furthermore, Adelson teaches using cumulative information of each frame to construct a lattice comprising the larger scene (mosaicing background), and also teaches displaying any portion of a scene that is desired to show, and also teaches extrapolating motion from other frames (Col.14, Figs.7A & 7B), implicitly teaching intra-scene motion analysis since each scene comprises of at least one frame, and conversely, each frame comprises a portion of the scene. Yeo discloses dividing video sequence into equal length segments, denoting the first frame of each segment as its key frame (Col.1: lines 34-38), and also teaches classifying a long video sequence into story units (Col.1: lines 47-50) based on content, using temporal segmentation of video based on scene change detection. Yeo's keyframes teach the concept of detecting scene transitions between video sequences. Adelson teaches background mosaicing, and also teaches a foreground object that moves over the stationary background mosaic, and also teaches displaying portions of the scene that is desired to show, as explained above, and also teaches that while slow movements can be encoded as warps of a single layer, faster moving objects should be split into different layers (Col.15). It would have been obvious to use intra-scene motion analysis to split the

video information into layers, and make the static background mosaic of Adelson the keyframe, since this would provide a more efficient encoding of images. As per the Shibata reference, Shibata was used in the office action to teach the steps of segmenting a video sequence, wherein individual video frames formed the smallest unit of any segment, and assuming a collection of basic segments as the initial layer, wherein a basic segment is a collection of video frames having the same vector expressions, and creating new layers by adding a segment to the previously processed layer, thus teaching a method for providing background mosaic, and intra-scene motion analysis. Though Shibata provides a textual description of the underlying video scene so that the video may be processed within the context of a video editing environment, the script is derived from segmentation of the video frames, and hence such concepts are implicitly taught therein. However, it is evident that the concepts that are being cited in Shibata are already taught by the Adelson and Yeo references.

As per appellants argument, regarding claims 4 and 24, that there in no teaching in the references to a network of mosaics, it is noted that the claims cite "one of a two-dimensional mosaic, a three-dimensional mosaic, and a network of mosaics", and Adelson teaches mosaics.

As per appellants arguments, regarding claims 5-8, that Adelson teaches away from subband encoding, it is noted that it is nevertheless disclosed therein. As per the argument that Jaillon's use of image pyramids is inappropriate to the invention, it is noted that the Laplacian

pyramids as taught by Jaillon may be used to merge images in accordance with their features, a concept which could be used in the mosaicing of the background layer, as taught by Adelson.

As per claims 13 and 14, Zhang teaches calculating the differences between consecutive video frames based on the selected difference metric, and defining a cut if the values exceed a threshold value (Fig.5 & 5A).

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

Mark Zimmerman

MARK ZIMMERMAN
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER ~~2700~~ 2600

MP
Mano Padmanabhan
October 20, 2000

Conf
km

Thomason, Moser & Patterson LLP
Attorneys at Law
595 Shrewsbury Avenue, 1st Floor
Shrewsbury, New Jersey 07702